Закон больших чисел

[массовые явления и закон больших чисел - закон больших чисел в форме Чебышева – пример - теорема Чебышева - теорема Бернулли - центральная предельная теорема Ляпунова]

- *Закон больших чисел* в теории вероятностей утверждает, что эмпирическое среднее (среднее арифметическое) достаточно большой конечной выборки из фиксированного распределения близко к теоретическому среднему – математическому ожиданию этого распределения.

Общий смысл закона больших чисел — совместное действие большого числа одинаковых и независимых случайных факторов приводит к результату, в пределе не зависящему от случая.

В теории вероятностей большое значение имеет установление закономерностей, происходящих с вероятностями близкими единице. Особую роль играют закономерности, возникающие в результате наложения большого числа независимых или слабо зависимых случайных факторов.

- Рассмотрим случайные величины x_n , являющиеся некоторыми заданными симметрическими функциями от первых n величин последовательности $\zeta_n = f_n(\xi_1, \xi_2, \dots, \xi_n)$. Если существует последовательность $a_1, a_n, \dots, a_n, \dots$, такая что при любом $\varepsilon > 0$

$$\lim_{n\to\infty} P\{|\xi_n - a_n| < \varepsilon\} = 1$$

то последовательность подчиняется закону больших чисел с заданными функциями f_n.

Часто эта функция – *среднее арифметическое* $\frac{1}{n}\sum_{i=1}^{n}\xi_i$ *величин* ξ_i,

$a_i = M\xi_i$.

Слабый закон больших чисел.

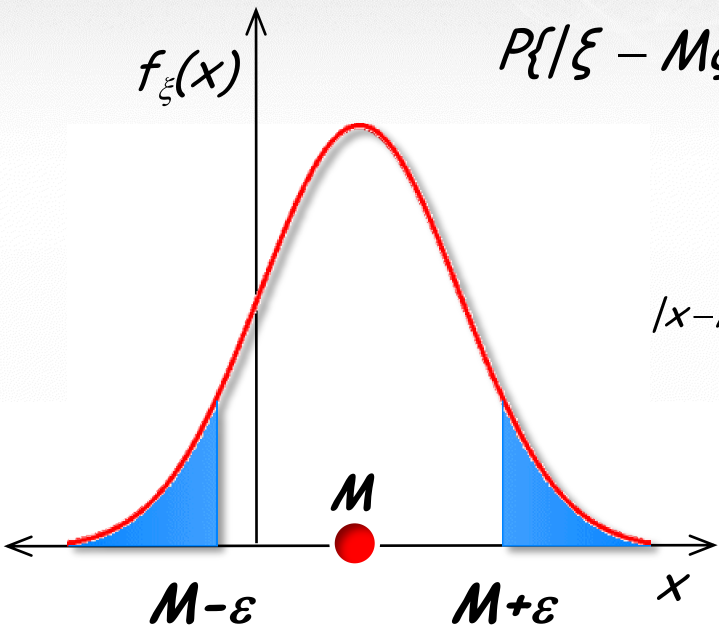
Закон больших чисел в форме Чебышева

- Для любой случайной величины ξ, имеющей конечную дисперсию, при каждом $\varepsilon > 0$ имеет место неравенство

$$P\{|\xi - M\xi| \geq \varepsilon\} \leq \frac{D\xi}{\varepsilon^2}$$

Неравенство Чебышева.

Доказательство



$$P\{|\xi - M\xi\,) \geq \varepsilon\} = \int\limits_{|x - M\xi| \geq \varepsilon} dF_\xi(x)$$

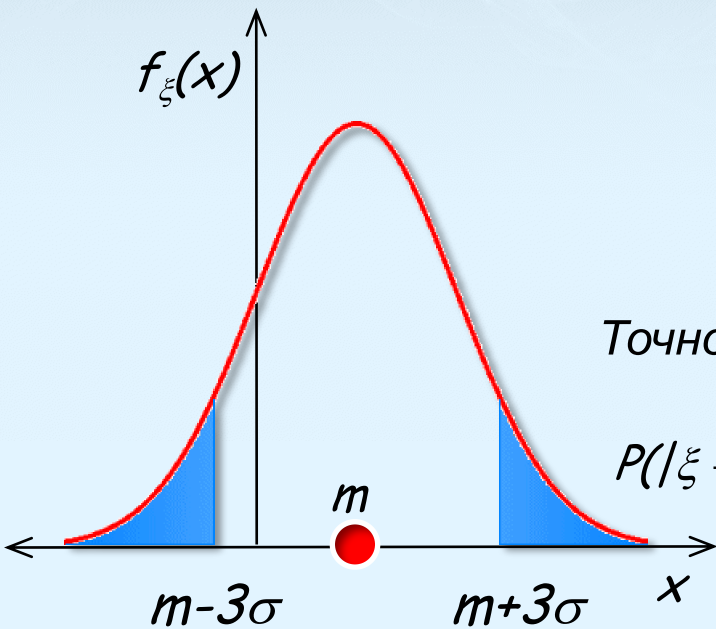
$$\frac{|x - M\xi|}{\varepsilon} \geq 1$$

$$\int\limits_{|x - M\xi| \geq \varepsilon} dF_\xi(x) \leq \frac{1}{\varepsilon^2} \int\limits_{|x - M\xi| \geq \varepsilon} (x - M\xi)^2\, dF_\xi(x)$$

$$\int\limits_{|x - M\xi| \geq \varepsilon} dF_\xi(x) \leq \frac{1}{\varepsilon^2} \int\limits_{-\infty}^{+\infty} (x - M\xi)^2\, dF_\xi(x) = \frac{D\xi}{\varepsilon^2}$$

$\square$

@ *Для случайной величины, распределенной по нормальному закону, найти оценку того, что случайная величина не попадет в интервал (m - 3σ ; m + 3σ)*



Решение

$$P(|\xi - m| \geq 3\sigma) \leq \frac{D\xi}{\varepsilon^2} = \frac{\sigma^2}{(3\sigma)^2} = \frac{1}{9}$$

Это самая верхняя граница для оценки

Точное решение

$$P(|\xi - m| \geq 3\sigma) = 1 - \frac{1}{\sqrt{2\pi}} \int_{\frac{m-3\sigma}{\sigma}}^{\frac{m+3\sigma}{\sigma}} e^{-\frac{t^2}{2}} dt = 1 - 2\Phi(3) = 0.0027$$

Если $\xi_1, \xi_2, \dots, \xi_n$ – *последовательность попарно независимых с.в., имеющих конечные дисперсии, ограниченные одной и той же постоянной* $D\xi_1 < C, D\xi_2 < C, \dots$, *то каково бы не было постоянное* $\varepsilon > 0$, *справедливо следующее:*

$$\lim_{n\to\infty} P\left\{\left|\frac{\sum_{k=1}^{n}\xi_k}{n} - \frac{\sum_{k=1}^{n}M\xi_k}{n}\right| < \varepsilon\right\} = 1$$

Доказательство

$$\bar{\xi} = \frac{\sum_{k=1}^{n}\xi_k}{n}, \quad M\bar{\xi} = \frac{\sum_{k=1}^{n}M\xi_k}{n}$$

На основании неравенства Чебышева

$$P\{|\bar{\xi} - M\bar{\xi}| < \varepsilon\} > 1 - \frac{D\bar{\xi}}{\varepsilon^2}$$

$$D\bar{\xi} = \frac{\sum_{k=1}^{n}D\xi_k}{n^2} < \frac{C}{n}$$

$$P\{|\bar{\xi} - M\bar{\xi}| < \varepsilon\} > 1 - \frac{C}{n\varepsilon^2}$$

Перейдем к пределу и получим доказательство

Центральная предельная теорема Ляпунова относится к законам распределения случайных величин и устанавливает условия, при которых возникает нормальный закон распределения.

- Пусть ξ_1, ξ_2, ... - независимые и одинаково распределенные случайные величины с конечной и ненулевой дисперсией $0 < D\xi_1 < \infty$.
Тогда имеет место слабая сходимость:

$$\frac{\xi_1 + \xi_2 + \ldots + \xi_n - nM\xi_1}{\sqrt{nD\xi_1}} \Rightarrow N(0,1)$$

последовательности «центрированных и нормированных» сумм случайных величин к стандартному нормальному распределению.

Если X представляет собой сумму очень большого числа взаимно независимых случайных величин X_1, X_2, …., X_n, влияние каждой из которых на всю сумму ничтожно мало, то величина X имеет распределение вероятности, близкое нормальному.